International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 3, 2025

The Use of Deep Learning in Enhancing the Accuracy of Weather Prediction Model

Ms.M.Kamala (Associate professor)
Department of CSE
CMR College of engineering and technology
mkamala@cmrcet.ac.in

Golla Saidulu (Assistant professor)
Department of CSE
CMR College of engineering and technology
gsaidulu@cmrcet.ac.in

Dr.G.Ravi Kumar (Associate professor)
Department of CSE
CMR College of engineering and technology
gravikumar@cmrcet.ac.in

Vempati Venkata Sai Charan Reddy
Department of CSE
CMR College of engineering and technology
20h51a0578@cmrcet.ac.in

Abstract: Weather predictions are important in several sectors like farming, transport, flood control and safety. But should it not go further under complex non-linear atmospheric data and dynamic weather patterns? Traditional numerical weather prediction models are reliable but there are often challenges when it comes to dealing with such complexities. Due to the advent of artificial intelligence, there has been the introduction of deep learning as an effective way of increasing the accuracy of weather prediction. In this paper, how deep learning processes like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks are used in improving both short-term and long-term weather forecasting is discussed. The models can process bulk of the time-series and spatial data to reveal concealed trends which are not identified by conventional models. By analyzing current studies and test outcomes, the paper demonstrates the importance of deep learning models as they minimize the estimation error and learn to adapt to real-time data much more easily than many of their alternatives. These findings answer in an affirmative way to the possible potential of deep learning that can complement or even outdo the traditional methods in providing more accurate and timely weather predictions.

Keywords: Weather prediction, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM).

I. INTRODUCTION

Forecasting is especially relevant in farming, transport, emergency planning and all everyday decision-making processes. Correct predictions can save crops, early weather warnings of natural calamities are made by governments and people were made ready. Nonetheless, it is a complicated job to forecast weather and according to the dynamics and chaotic conditions of atmospheric conditions, it is not easy. Conventional forecasting systems may not be equipped to deal with the sheer number and volatility of real-time weather data and this makes them not capable of providing timely and accurate forecast.

Recent years brought a breakthrough in the field of artificial intelligence (AI) and especially deep learning, which provides new opportunities to enhance the accuracy of weather forecasting. Deep learning, notably neural networks are structured in such a way that patterns and features are learned in big datasets. It makes them very suitable in handling Satellite, Radar measurements, temperature, humidity, wind, and so on. In contrast to rule-based systems, deep learning algorithms are more inclusive than models and may learn new data and identify unknown relationships that traditional models could overlook.

Among the greatest strengths of deep learning in weather prediction, one can notice the handling of non-linear relationships, as well as time series information. Recurrent Neural Networks (RNNs) and their superior types such as Long Short-Term Memory (LSTM) networks are the most popular in approximating time-based data. Such models have the capacity to know how the present weather conditions are affected by previous conditions which enhances both long range and short range forecasts. Such a time awareness is important in forecasting Rainfalls, temperature variation and development of the storms.

Several studies have shown that deep learning models outperform traditional numerical weather prediction methods in specific tasks such as rainfall forecasting, fog detection, and storm intensity estimation. While deep learning does not completely replace physics-based models, it serves as a powerful complement. By integrating deep learning outputs with conventional models, hybrid systems achieve more balanced and reliable forecasts.

In this paper, we explore how deep learning techniques enhance the accuracy of weather prediction systems. We examine commonly used architectures such as CNNs, RNNs, LSTMs, and hybrid models. We also discuss the challenges in data collection, model training, and validation, and suggest potential improvements for future applications. The aim is to demonstrate the practical value of deep learning in modern meteorology and how it contributes to building smarter, faster, and more accurate weather prediction systems.

II. LITERATURE SURVEY

Frnda et al. [1] investigated the possibility of deep learning in enhancing the performance of predictions of ECMWF (European Centre for Medium-Range Weather Forecasts) particularly on short term weather forecasting of up to three days. They have come up with a calibration model based on a neural network to rectify mistakes in temperature and precipitation prediction. Their methodology is that they mix raw ECMWF sources filters with other environmental parameters based on real weather data provided by various cities. Their model results demonstrated that it substantially enhanced the accuracy of their forecast, thus it is at par with high-precision local models.

Abdulla et al. [2] targeted improving weather prediction regarding adaptive deep learning types that rely on Long Short-Term Memory (LSTM). To know the most accurate model of LSTM, they experimented with various types of LSTM as well as employing the adaptive learning by frequently updating the model with new data. They demonstrated through their experiments how the bidirectional LSTM model when used with the adaptive learning, lower prediction errors than the baseline model by 45%. They also discovered that univariate models (with less features) were more effective in its time and memory efficiency.

According to Kumar et al. [3], weather forecasting became particularly significant in such areas of economy as agriculture and transportation. In their research, Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs) are applied as deep learning models analyzing weather patterns. They demonstrated that the deep learning approach was better than the traditional approaches, particularly the analysis of complex and dynamic data regarding the climate. Their outputs pointed to the fact that RNNs were especially effective in modelling the time-based weather patterns and providing accurate forecasts.

Salman et al. [4] investigated deep learning model (RNN, Conditional Restricted Boltzmann Machine (CRBM) and Convolutional Networks (CN) to predict weather. They compared the accuracy of each model using weather records of weather station in Indonesia and global datasets such as ENSO. Their search identified that deep learning has the potential to identify concealed trends in giant weather data and make better predictions of all kinds of applications such as agriculture and aviation. They tested their models with mathematical accuracy quantifiers and indicated a good potential.

The article by Yadav et al. [5] covered the issue of predicting weather conditions in remote areas where data was usually scanty or absent. They outlined the development of a deep learning model that may learn using the well-documented spots and then forecast weather in poorly-documented spots. They had an accuracy of 83 percent in the new method and it was evident that traditional methods of forecasting were much worse. The strategy demonstrates that, despite a lack of data in some regions, AI models can become a solution to the problem by reducing the gap and delivering accurate predictions, making underserved populations prepared to extreme weather conditions.

Chen et al. [6] has given a detailed summary of the utilization of machine learning in weather and climate prediction. They surveyed over 20 available ML approaches and pointed out eight of them that are particularly effective to expect. According to their survey, ML is already quite effective in the context of short-term forecasting, whereas its implementation concerning long-term climate forecast has not yet become unconstrained with regard to the quality of data and the complexity of models. They pointed out the importance to intensify the interaction between weather and climate models and stimulated additional interdisciplinary studies in this matter.

Shah et al. [7] have been studying rainfall prediction improvements based on statistical methods and models of machine learning. They compared such models as ARIMA, Neural Networks, and Random Forest based on daily weather data. Out of these, Random Forest model had the highest prediction to seasonal rainfall. Their study was especially helpful to farmers and researchers who require precise data in their planning and decision-making to increase their accuracy in predicting significant aspects of weather such as rain and temperature among others and thus proving that AI models can increase accuracy in predicting these critical aspects of weather.

III. PROPOSED METHOD

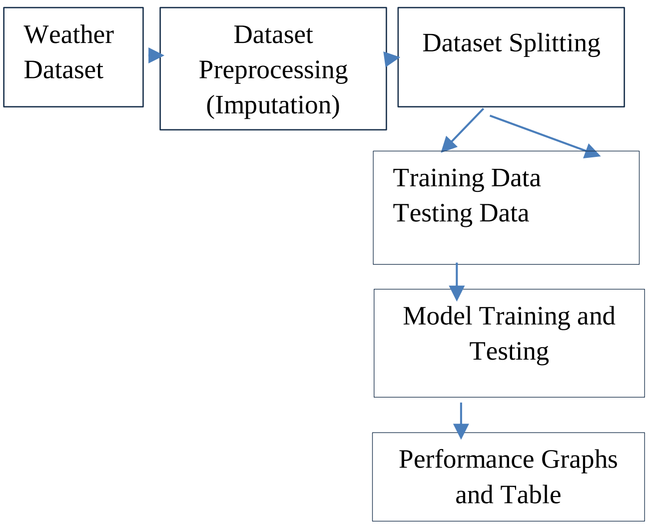

Fig.3.1 Block diagram of proposed method

The suggested approach will achieve higher precision in forecasting the weather conditions with the aid of deep learning practices. It combines the process of data preprocessing, cleaning, and training of several machine learning models, to enhance the prediction results. They are as follows:

3.1 Data preparation/Augmentation and concatenation: Raw weather data would be gathered in 4 different weeks. In order to have a perfect analysis, all these datasets are maintained into one unison dataset via the concat function in Python. The last data set has a term of labels of attributes on the top and subsequent rows having the meaning of values.

3.2 Data Preprocessing:

The reason why preprocessing should be an essential part of data analysis is that it allows cleaning the dataset and enhances the quality of data fed into the models. It includes:

- Missing Value Imputation: The method of estimating the missing values should be used as a K-Nearest Neighbors (KNN) algorithm is used to fill the missing data slot with the estimate of the missing data based on a similar case.

- Data Standardization: the fuzzy C-means clustering will be utilized in order to normalize the given dataset and to make it more standard and uniform.

- Filling in the gaps: The missing values are filled, and all the categorical data is converted into numerical one, and it can be adapted to the needs of machine learning.

3.3 Model training and evaluation:

The trained data is segregated into training and test data. A number of deep learning models are subsequently trained upon this data and tested:

Support Vector Machine (SVM): It undertakes baseline forecasting, which reflects at minimal error between observed and forecasted weather.

Convolutional Neural Network (CNN): Analyzes spatial patterns to evaluate the data and show high accuracy in forecasting the air quality indices.

Recurrent Neural Network (RNN): Quantifies the temporal nature of the data, and it can obtain reliable weather prediction with low deviation.

International Journal of Information Technology & Computer Engineering

ISSN 2347-3657

Volume 13, Issue 3, 2025

Story time Long short-term memory (LSTM): A more sophisticated version of RNN to deal with long-term dependencies, which is also more accurate in predicting.

3.4 Performance Analysis:

The models performance is measured by looking at the difference between the actual observations and the predictions. To determine the accuracy of the models, graphical depictions (line plots, performance curves) are drawn to represent and quantitatively assess the level of accuracy.

4 RESULT

Proposed method is implemented using python programming and Jupyter notebook.We have different datasets week-wise so below we used 'concat' function to concatenate the multiple datasets into single dataset.

```
#now merging all datasets from data1 to data6 as single dataset
dataset = pd.concat([data1, data2, data3, data4])
dataset
```

	Time	AQI-IN	AQI-US	PM25	PM10	PM1	Temp(cel)	Hum	Noise	TVOC(ppm)	...	AQI-IN(F)	AQI-IN(s)	CI	VI	particle count(0.3)	particle count(0.5)
0	07-03-2024 23:45	121.778	156.511	66.533	87.356	62.867	24.102	97.753	46.978	0.006	...	121.778	121.778	10.000	10.0	20119.533	2465.733
1	07-03-2024 23:30	126.818	157.341	68.045	90.318	64.068	24.170	97.611	47.705	0.006	...	126.818	126.818	10.000	10.0	20450.250	2539.296
2	07-03-2024 23:15	111.756	154.822	63.467	82.756	60.111	24.238	97.060	47.067	0.006	...	111.756	111.756	9.889	10.0	19181.467	2364.000
3	07-03-2024 23:00	119.659	156.159	65.886	87.136	62.364	23.941	97.611	47.409	0.006	...	119.659	119.659	10.000	10.0	19893.682	2455.364
	07-																

Fig:4.1 Final Dataset after concatenation of 4-week Dataset

We have 4 different weeks dataset. To work on whole data , we have concatenated all4 datasets to a single dataset. Above dataset contains first row as attributes and remaining rows contains the values.

```
#KNN Computation
from sklearn.impute import KNNImputer
# Display missing values in the dataset
print("Missing values in each column:")
print(dataset.isnull().sum())

# Select features of interest for imputation
columns_of_interest = ['PM25', 'PM10', 'Temp(cel)', 'Hum', 'Noise']
dataset_selected = dataset[columns_of_interest]

# Apply KNN Imputer
imputer = KNNImputer(n_neighbors=5)  # K=5
dataset_imputed_knn = pd.DataFrame(imputer.fit_transform(dataset_selected)

# Display the imputed dataset
print("Data after KNN imputation:")
print(dataset_imputed_knn.head())
```

```
Missing values in each column:
Time                 0
AQI-IN               0
AQI-US               0
PM25                 0
```

Fig:4.2 KNN Computation for preprocessing

KNN computation is applied as preprocessing step. Preprocessing helps to remove unwanted data by operating on missing values etc.

```
#Fuzzy C means
from fcmeans import FCM

# Select relevant columns for imputation
columns_of_interest = ['PM25', 'PM10', 'Temp(cel)', 'Hum', 'Noise']
dataset_selected = dataset[columns_of_interest]

# Normalize the data (Fuzzy C-Means requires normalized input)
scaler = MinMaxScaler()
dataset_normalized = pd.DataFrame(scaler.fit_transform(dataset_selected), columns=columns_of_interest)

# Separate rows with missing values
missing_indices = dataset_normalized[dataset_normalized.isnull().any(axis=1)].index
non_missing_data = dataset_normalized.dropna()

# Apply Fuzzy C-Means clustering
fcm = FCM(n_clusters=3, random_state=42)  # Number of clusters = 3
fcm.fit(non_missing_data.values)

# Get cluster centroids and memberships
centroids = fcm.centers
memberships = fcm.u
```

Fig:4.3 Applied fuzzy C-means

Applied fuzzy c-means algorithm as preprocessing step to make data in more standard format.

Dataset After Cleaning & Processing

	AQI-IN	AQI-US	PM25	PM10	PM1	Temp(cel)	Hum	Noise	TVOC(ppm)	CO(ppm)	...	particle count(1.0)	particle count(3.0)	particle count(5.0)	particle count(10.0)	year_value
0	121.778	156.511	66.533	87.356	62.867	24.102	97.753	46.978	0.006	0.471	...	352.556	145.800	0.000	0.0	2024
1	126.818	157.341	68.045	90.318	64.068	24.170	97.611	47.705	0.006	0.431	...	365.591	154.114	0.000	0.0	2024
2	111.756	154.822	63.467	82.756	60.111	24.238	97.060	47.067	0.006	0.395	...	331.422	135.556	0.000	0.0	2024
3	119.659	156.159	65.886	87.136	62.364	23.941	97.611	47.409	0.006	0.357	...	349.591	145.932	0.000	0.0	2024
4	117.705	155.932	65.273	85.182	61.818	23.961	96.973	46.727	0.006	0.379	...	339.227	137.932	0.068	0.0	2024
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1919	55.182	95.205	33.068	40.682	31.659	34.161	55.245	50.045	0.008	0.240	...	148.477	58.364	0.000	0.0	2024
1920	57.311	98.156	34.400	42.222	32.822	35.022	55.109	49.978	0.008	0.240	...	148.822	59.711	0.000	0.0	2024
1923	59.400	101.422	35.667	43.622	34.067	34.798	55.971	50.156	0.008	0.211	...	158.156	61.956	0.000	0.0	2024
1924	62.267	105.556	37.311	46.333	35.689	34.469	57.191	50.489	0.008	0.205	...	170.333	68.889	0.000	0.0	2024
1925	62.933	106.667	37.778	46.267	36.089	34.707	56.809	49.800	0.009	0.182	...	166.844	64.933	0.000	0.0	2024

Fig:4.4 Dataset After Cleaning and Processing

After preprocessing data obtained is shown above which contains all the rows and columns are filled and the categorial data if any is converted to numerical data.

```
True Air Quality Index : 65.48899999999999 Predicted Air Quality Index : 70.86228286844056
True Air Quality Index : 68.227 Predicted Air Quality Index : 73.17850822879552
True Air Quality Index : 74.0 Predicted Air Quality Index : 85.33606949092847
True Air Quality Index : 145.122 Predicted Air Quality Index : 141.7507231255193
True Air Quality Index : 47.822 Predicted Air Quality Index : 48.28874632734555
True Air Quality Index : 51.288999999999994 Predicted Air Quality Index : 61.903775599562124
True Air Quality Index : 97.26700000000001 Predicted Air Quality Index : 116.87018809463679
True Air Quality Index : 60.599999999999994 Predicted Air Quality Index : 66.73132242468212
True Air Quality Index : 101.281 Predicted Air Quality Index : 115.32917719547675
True Air Quality Index : 52.977999999999994 Predicted Air Quality Index : 58.136264861047856
```

Fig:4.5 Actual weather quality values and SVM model Predicted Values

For given test data , actual WEATHER FORECASTING value and predicted WEATHER FORECASTING value by SVM are shown. The obtained value difference is very less.

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 3, 2025

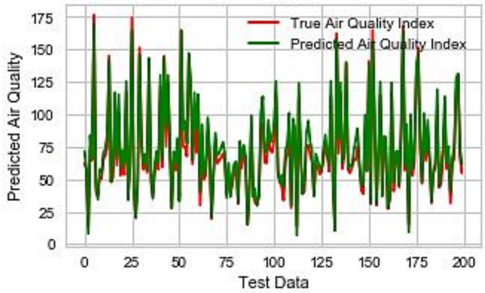

Fig:4.6 CNN model Performance

Proposed CNN algorithm performance is shown in above figure which shows the true air quality index and predicted air quality index.

```
True Air Quality Index : 50.356 Predicted Air Quality Index : 50.356
True Air Quality Index : 65.48899999999999 Predicted Air Quality Index : 65.48899999999999
True Air Quality Index : 68.227 Predicted Air Quality Index : 68.227
True Air Quality Index : 74.0 Predicted Air Quality Index : 74.0
True Air Quality Index : 145.122 Predicted Air Quality Index : 145.122
True Air Quality Index : 47.822 Predicted Air Quality Index : 47.822
True Air Quality Index : 51.288999999999994 Predicted Air Quality Index : 51.288999999999994
True Air Quality Index : 97.26700000000001 Predicted Air Quality Index : 97.26700000000001
True Air Quality Index : 60.599999999999994 Predicted Air Quality Index : 60.599999999999994
True Air Quality Index : 101.281 Predicted Air Quality Index : 107.98679021107219
True Air Quality Index : 52.977999999999994 Predicted Air Quality Index : 52.977999999999994
```

Fig:4.7 Actual WEATHER FORECASTING values and RNN model Predicted Values

For given test data , actual WEATHER FORECASTING value and predicted WEATHER FORECASTING value by RNN are shown. The obtained value difference is very less.

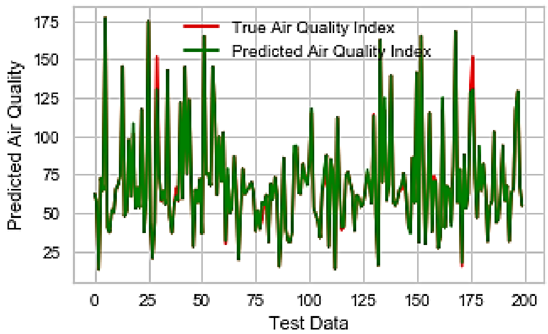

Fig:4.8 RNN Model Performance

RNN algorithm performance is shown in graphical format which represents true air quality index and predicted air quality index.

```
True Air Quality Index : 51.711000000000001 Predicted Air Quality Index : 51.95772050269212
True Air Quality Index : 50.356 Predicted Air Quality Index : 51.31598299121271
True Air Quality Index : 65.48899999999999 Predicted Air Quality Index : 64.23988559817124
True Air Quality Index : 68.227 Predicted Air Quality Index : 68.84833048053471
True Air Quality Index : 74.0 Predicted Air Quality Index : 75.81356359262357
True Air Quality Index : 145.122 Predicted Air Quality Index : 140.6925441637727
True Air Quality Index : 47.822 Predicted Air Quality Index : 44.87375798395868
True Air Quality Index : 51.288999999999994 Predicted Air Quality Index : 51.241792679771955
True Air Quality Index : 97.26700000000001 Predicted Air Quality Index : 102.4212482399891
True Air Quality Index : 60.599999999999994 Predicted Air Quality Index : 61.27849587913109
True Air Quality Index : 101.281 Predicted Air Quality Index : 109.05196644771384
True Air Quality Index : 52.977999999999994 Predicted Air Quality Index : 51.29900818945228
```

Fig:4.9 Actual WEATHER FORECASTING values and LSTM model Predicted Values

For given test data , actual WEATHER FORECASTING value and predicted WEATHER FORECASTING value by LSTM are shown. The obtained value difference is very less.

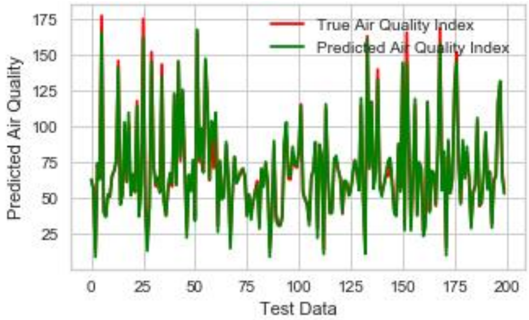

Fig:4.10 LSTM Model Performance

LSTM algorithm performance is shown in graphical format which represents true air quality index and predicted air quality index.

5 CONCLUSION

This study concludes that deep learning techniques offer substantial improvements in the accuracy and efficiency of weather prediction models. By learning complex patterns from historical and real-time weather data, models like CNNs and LSTMs are capable of delivering more precise short-term and long-term forecasts. These advancements are particularly beneficial in scenarios where traditional numerical models fall short, such as in remote regions with limited data or during rapidly changing weather events. Deep learning models not only enhance the reliability of weather forecasts but also allow for faster processing and easier adaptation to new information. As technology and data availability continue to grow, the integration of deep learning into mainstream meteorological systems will play a crucial role in improving public preparedness and minimizing the impact of adverse weather conditions. Future research should focus on hybrid models, improved data quality, and real-time deployment to unlock the full potential of AI-driven weather forecasting.

REFERENCES

[1] Frnda, J., Durica, M., Rozhon, J., Vojtekova, M., Nedoma, J., & Martinek, R. (2022). ECMWF short-term prediction accuracy improvement by deep learning. *Scientific Reports*, *12*(1), 7898.

[2] Abdulla, N., Demirci, M., & Ozdemir, S. (2022). Design and evaluation of adaptive deep learning models for weather forecasting. *Engineering Applications of Artificial Intelligence*, *116*, 105440.

[3] Kumar, N., Dash, Y., Abraham, A., Ahmad, S. I., & Saifullah, S. (2023). Enhancing weather forecast accuracy using deep learning models. In *International Conference on Soft Computing and Pattern Recognition* (pp. 527–537). Springer Nature Switzerland.

[4] Salman, A. G., Kanigoro, B., & Heryadi, Y. (2015). Weather forecasting using deep learning techniques. In *2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*.

[5] Yadav, K., Malviya, S., & Tiwari, A. K. (2025). Improving weather forecasting in remote regions through machine learning. *Atmosphere*, *16*(5), 587.

[6] Chen, L., Han, B., Wang, X., Zhao, J., Yang, W., & Yang, Z. (2023). Machine learning methods in weather and climate applications: A survey. *Applied Sciences*, *13*(21), 12019.

[7] Shah, U., Garg, S., Sisodiya, N., Dube, N., & Sharma, S. (2018). Rainfall prediction: Accuracy enhancement using machine learning and forecasting techniques. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 776–782). IEEE.